AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY 100:191–206 (1996)

Estimating Sexual Dimorphism by Method-of-Moments

STEVEN C. JOSEPHSON, KENNETH E. JUELL, AND
ALAN R. ROGERS
Department of Anthropology, University of Utah, Salt Lake City, Utah 84112

KEY WORDS Sexual dimorphism, Method of moments, Coefficient of variation method, Mean method

ABSTRACT Estimating the degree of sexual dimorphism is difficult in fossil species because most specimens lack indicators of sex. We present a procedure that estimates sexual dimorphism in samples of unknown sex using method-of-moments. We assume that the distribution of a metric trait is composed of two underlying normal distributions, one for males and one for females. We use three moments around the mean of the combined-sex distribution to estimate the means and the common standard deviation of the two underlying distributions. This procedure has advantages over previous methods: it is relatively simple to use, specimens need not be assigned to sex a priori, no reference to living species analogs is required, and the method provides conservative estimates of dimorphism under a variety of conditions. The method performs best when the male and female distributions overlap minimally but also works well when overlap is substantial. Simulations indicate that this relatively simple method is more accurate and reliable than previous methods for estimating dimorphism.

Several lines of evidence suggest that sexual dimorphism is related to mating systems and associated behavior, such as intrasexual competition and male parental investment among primates and other animals (Andersson, 1994; Clutton-Brock and Harvey, 1977, 1978; Jarman, 1983; Plavcan and van Schaik, 1992a,b; Ralls, 1977; Richard, 1992). Estimates of sexual dimorphism have been used to infer similar behaviors in fossil primate species (e.g., Andrews, 1983; Fleagle et al., 1980; Kay 1982a,b; Leutenegger and Shell, 1987; McHenry, 1991). The contribution of sexual dimorphism to intraspecific morphological variability in extant and fossil species has received considerable attention (e.g., Fleagle et al., 1980; Godfrey et al., 1993; Kay, 1982a,b; Kay et al., 1988; Krishtalka et al., 1990; Pickford and Chiarelli, 1986; Plavcan and Kay, 1988; Plavcan and van Schaik, 1992a,b; Plavcan, 1994). Dimorphism has also figured prominently in discussions regarding the accurate assignment of fossil specimens to species (Cope, 1993; Kelley, 1993; Plavcan, 1993; Teaford et al., 1993), such as recent debates concerning the hominid specimens placed within *Australopithecus afarensis* (Foley and Lee, 1989; Johanson and White, 1979; Kimbel and White, 1988; McHenry, 1991; Stern and Susman, 1983; Zihlman, 1985) and *Homo habilis* (Groves, 1989; Lieberman et al., 1988; Miller, 1991; Rightmire, 1993; Tobias, 1978, 1991; Walker, 1981; Wood, 1985, 1991, 1992). These debates in part have involved disagreements over methods used to estimate dimorphism and would benefit from an unbiased, reliable method for estimating sexual dimorphism in species known only through fossils.

Received April 18, 1995; accepted December 31, 1995.

Address reprint requests to Steven C. Josephson, 117 Stewart Building, Department of Anthropology, University of Utah, UT 84112.

Primate and hominid fossils are often fragmentary, making it difficult or impossible to assign sex to specimens. In the absence of a reliable sex marker, some studies have assigned sex to specimens by assuming that larger ones are males and smaller ones females. This procedure is tautological because it presupposes sexual dimorphism in order to estimate it. Techniques that assume that the distributions of males and females are nonoverlapping are similarly difficult to support (see Godfrey et al., 1993; Plavcan, 1994). Even in moderately dimorphic species, male and female distributions can overlap considerably, and assigning sex by size can overestimate dimorphism (see discussion in Godfrey et al., 1993; below).

Also, different traits within a species may display different degrees of dimorphism (Oxnard, 1983). Canine teeth, for example, are sometimes highly dimorphic in species where other craniodental traits show little or no sexual dimorphism (e.g., Fleagle et al., 1980; Gingerich, 1981; Kay, 1982a; Kelley and Xu, 1991; Krishtalka et al., 1990; Martin, 1991; Pickford, 1986). Conversely, some hominid taxa may be moderately to highly dimorphic in body size (Jungers, 1990; McHenry, 1991) but only slightly dimorphic in canines (see Leutennegger and Shell, 1987). We need a technique that can reliably determine degrees of dimorphism in different traits.

We present a technique that has assumptions similar to previous ones that correlate combined-sex sample variability with sexual dimorphism (Fleagle et al., 1980; Godfrey et al., 1993; Kay, 1982a,b; Plavcan, 1994). Unlike previous techniques, no a priori sex assignment is necessary, no extant species analogs are required to interpret results, and any number of metric traits may be analyzed independently of each other.

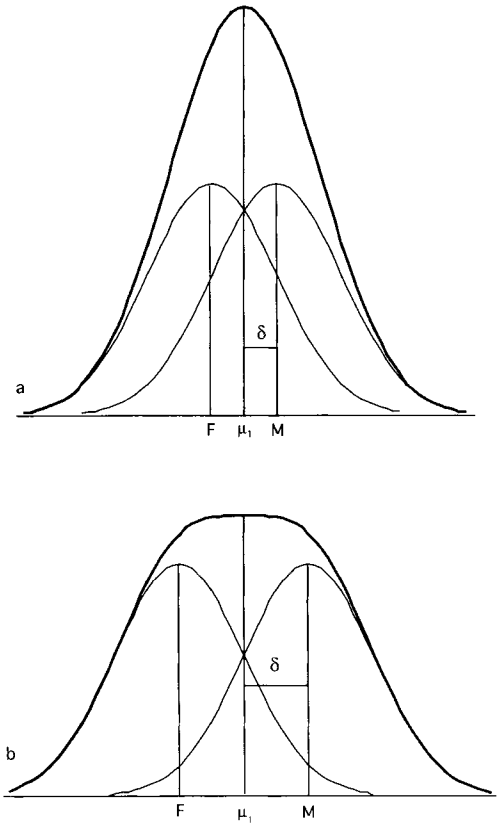

Fig. 1. The relationship between the mean of the male distribution (M), the mean of the female distribution (F), and the distance between M or F and m_1 (δ) when **(a)** $\delta = \frac{1}{2}\sigma$ and **(b)** $\delta = \sigma$.

METHODS

Our method-of-moments (MoM) technique assumes that each subject is equally and independently likely to be a male or a female. Each male is drawn independently from a normal distribution with a mean of $\mu_1 + \delta$ and a variance of σ^2, and where μ_1 is the mean of the combined-sex distribution and δ is the distance between the mean for the male distribution (M) and the combined-sex mean μ_1 (see Fig. 1a). Each female is drawn independently from a normal distribution with a mean of $\mu_1 - \delta$, and a variance of σ^2. The equations that describe such distributions are well known and have been used in a number of applications (see Everett and Hand, 1981), but not to estimate sexual dimorphism.

We use the equations from Johnson and Kotz (1970: 89) for expressing the theoretical moments about the origin for the combined-sex distribution in terms of the within-sex distributions. While we begin with data which are centered about the mean μ_1, we will subtract off the mean from each value

in a z-transform, after which the data center about the origin m_1.[1]

We refer to the mean of the larger sex as M (for "males") and the mean of the smaller sex as F (for "females") for convenience, but the solution is the same no matter which sex is larger. When a trait is dimorphic, the distance δ between M or F and μ_1 increases (Fig. 1b). As δ changes, so do the moments of the combined-sex distribution. Using equations from Johnson and Kotz (1970), we can express the variance of the within-sex distributions σ^2 and δ in terms of the moments about the origin of the combined-sex distribution (see Appendix for details):

$$\sigma^2 = -m_1^2 - \delta^2 + m_2 \qquad (1)$$

$$\delta = \left(-m_1^4 + \frac{3}{2}m_2^2 - \frac{1}{2}m_4\right)^{1/4}. \qquad (2)$$

Our estimate $\hat{\delta}$ of δ is obtained by equating the theoretical moments about the origin m_1, m_2, and m_4 with their estimates $\hat{m}_1$, $\hat{m}_2$, and $\hat{m}_4$. When the quantity in the parentheses in Equation 2 is negative, $\hat{\delta}$ is arbitrarily assigned a value of 0.

Equation 2 is simplified by transforming the original values of the sample distribution to z-scores. This is done by subtracting $\hat{\mu}_1$ from each observation and dividing it by $\sqrt{\hat{\mu}_2}$.

We define $\hat{\delta}^{(z)}$ as the value of $\hat{\delta}$ when the original values are transformed to z-scores. The z-transformed distribution now has $\hat{m}_1^{(z)} = 0$, and $\hat{m}_2^{(z)} = 1$, which makes Equation 2:

$$\hat{\delta}^{(z)} = \left(\frac{3}{2} - \frac{\hat{m}_4^{(z)}}{2}\right)^{1/4}. \qquad (3)$$

$\hat{\delta}^{(z)}$ then is transformed back to $\hat{\delta}$ by multiplying $\hat{\delta}^{(z)}$ by $\sqrt{\hat{\mu}_2}$ the standard deviation of the original sample distribution. Dimorphism measured on the normal scale, expressed as the male mean divided by the female mean, is $e^{2\delta}$.[2]

As a strictly numerical example, Table 1 contains data on the length of the left index fingers of the members of the Evolutionary Ecology Research Group at the University of Utah. There were nine males and eight females present at the meeting where we first presented this research, so the sex ratio is not perfectly balanced. The mean for the males M = 9.833 and the mean for females F = 8.786. Since we know the sex of these

[1]The observed moments about the mean are calculated from the combined-sex sample using the formula for the i^{th} moment about the mean

$$\hat{\mu}_i = \frac{1}{n}\sum_{j=1}^{N}(x_j - \mu_1)^i$$

where n is the sample size. The i^{th} moment about the origin is the expectation of the i^{th} power of a random variable, and is calculated using $\hat{m}_i = \frac{1}{n}\sum_{j=1}^{N} x_j^i$.

[2]Our method assumes that male and female values are each normally distributed, with identical variance within each sex. On the other hand, real characters are more often log-normally distributed, with a variance that is larger within the larger sex. Consequently, it is best to think of our formulas as applying to the logarithms of these "natural" character values. We will refer to the natural scale of measurement as the "log-normal scale," and to the scale within our formulas as the "normal scale." Actual data can be moved from the log-normal scale to the normal scale by ln-transforming each value, but understanding how mean and variance are related on the different scales is more complicated. The parameters of the normal scale of measurement can be related to means and variances on the log-normal scale using the properties of the log-normal probability distribution (Hastings and Peacock, 1974).

Let α and β denote the means for males and females on the log-normal scale, and let σ_m^2 and σ_f^2 denote the corresponding variances. Then:

$$\alpha = e^{(\mu + \delta + \sigma^2/2)}$$

$$\beta = e^{(\mu - \delta + \sigma^2/2)}$$

$$\sigma_m^2 = \alpha^2(e^{\sigma^2} - 1)$$

$$\sigma_f^2 = \beta^2(e^{\sigma^2} - 1).$$

On the log-normal scale, sexual dimorphism is usually measured by the ratio of mean male to mean female character values:

$$\alpha/\beta = e^{2\delta}.$$

Thus, δ and dimorphism can be estimated directly from the ratio of α and β. To obtain an estimate of σ^2, note that the squared coefficient of variation within each sex is

$$C = \sigma_m^2/\alpha^2 = \sigma_f^2/\beta^2 = e^{\sigma^2} - 1$$

To estimate σ^2 from data on the log-normal scale, we set C equal to the average squared CV within the two sexes and solve the resulting equation. μ can then be estimated as:

$$\hat{\mu} = \tfrac{1}{2}(\ln\alpha + \ln\beta - \hat{\sigma}^2).$$

TABLE 1. Measurements of the left index fingers of the members of the Evolutionary Ecology Research Group at the University of Utah

Males	Females
10.0	8.4
9.8	8.3
10.1	9.2
9.7	8.6
9.6	8.4
9.7	9.2
10.1	9.1
9.6	9.1
9.9	

TABLE 2. The length and breadth (in mm) of 13 third premolars of Australopithecus afarensis *(modified from Table 2, Johanson et al., 1983)*

Specimen	Maximum length	Minimum breadth
198-1	10.4	6.5
277-1	12.3	8.8
266-1	11.2	8.4
288-1i	10.0	7.1
311-1	10.5	6.9
128-23	9.7	6.5
207-13	10.6	7.5
400-1a	12.0	8.3
333w-1a,c	11.3	8.7
333w-46	11.1	9.0
333w-60, 32	12.6	9.5
333w-58	11.6	7.3
333-10	12.4	9.5

individuals, we can estimate the dimorphism simply as M/F = 1.119.

To estimate the dimorphism using the method-of-moments technique, we will assume that the sex of each individual is unknown. We begin by taking the natural log of each value. We then z-transform these values by subtracting off the mean ($\hat{\mu}_1 = 2.232$) and dividing by the standard deviation ($\sqrt{\hat{\mu}_2} = 0.065$). We then raise each of these values to the fourth power and find the average value ($\hat{m}_4^{(z)} = 1.944$). Substituting this value into Equation 3 gives $\hat{\delta}^{(z)} = 0.852$. To undo the z-transform, we multiply $\hat{\delta}^{(z)}$ by the ln-transformed sample standard deviation (0.065), which gives $\hat{\delta} = 0.056$. Our estimate of dimorphism in this sample when sex is unknown is $e^{2(0.056)} = 1.117$, which is very close to 1.119, the estimate of dimorphism when sex is known.

As an example using paleontological data, Table 2 contains a list of length and breadth measurements from 13 *A. afarensis* third premolars (Johanson et al., 1983). Multiplying the length by the width gives an estimate of the area in cross section of these teeth. The MoM estimate of $\hat{\delta}$ for this two-dimensional trait is 0.178, which makes the estimate of dimorphism 1.42.

Confidence intervals for $\hat{\delta}$

Estimates of dimorphism are only as good as the confidence we can place in them. Once we have a value for $\hat{\delta}$, we need to know how confident we can be about the accuracy of our estimate. To estimate confidence intervals for δ, we generate simulated samples with properties similar to our observed samples and simulate the stochastic processes which generate error. Certain parameters, such as our estimate of the variance σ^2 within the underlying male and female distributions and the sex ratio of the sample, will either be unknown or will vary between samples. We need to know how this variation affects $\hat{\delta}$.

We used a program to simulate sampling from the combined-sex distribution to evaluate this method. In order to simulate realistic data, we used values of δ and σ in approximate proportion to those seen in actual datasets: $\delta = \frac{1}{2}\ln(M/F)$ and $\sigma \approx$ within-sex coefficient of variation.[3] At each set of parameter values it performs the following operations 500 times: (1) Draw a random sample of n observations from a combined-sex distribution with a mean of μ_1, $\sigma = 1$, and a given value of δ.[4] (2) Use the sample to calculate $\hat{\delta}$. This yields 500 simulated values of $\hat{\delta}$ for each value of δ. The relative frequencies of the simulated values estimate the sampling deviations of our estimator. The closer $\hat{\delta}$ is to the value of δ, the better the estimate of dimorphism.

Note that the procedure does not strictly control some characteristics of the sample distribution, such as the sample sex ratio and sample variance within sexes. These parameters are allowed to vary much the same

[3]See Footnote 2.

[4]We lose no generality by restricting attention to particular values of μ_1 and σ because these parameters can always be made to equal 0 and 1 by a judicious choice of the scale of measurement.

way as they will when we sample fossil species via the paleontological record.

For each value of δ from 0.0 to 5.0 by increments of 0.5, we ran 500 trials of sample size n. The values for $\hat{\delta}$ in the trials were tabulated at each level of δ and used to derive an estimate of the central 95% of the observations, from the 0.025 quantile to the 0.975 quantile. We refer to this as the "spread" of $\hat{\delta}$.

A 95% confidence interval for δ comprises the set of hypotheses about δ that cannot be rejected at the 5% significance level. To estimate 95% confidence intervals of dimorphism in a paleontological sample, we need to know the range of values of δ which cannot be rejected at the 5% level. This is done by simulating 500 values of $\hat{\delta}$ at each of many levels of δ. Values of δ can be rejected with 5% confidence if the spread of the simulated values does not contain the value of $\hat{\delta}$ estimated in the paleontological sample.

With real data we cannot assume that $\sigma = 1$, therefore we estimate σ using Equation 4, substituting $\hat{\delta}$ for δ. Equation 4 is the simplified version of 1 and is derived by z-transforming the original sample. Since the transformed sample has a $\hat{m}_1^{(z)} = 0.0$, and a $\hat{m}_2^{(z)} = 1.0$, Equation 1 simplifies to:

$$\hat{\sigma} = \sqrt{(1 - (\hat{\delta}^{(z)})^2)}\sqrt{\hat{\mu}_2} \qquad (4)$$

where $\hat{\delta}^{(z)}$ is the value of $\hat{\delta}$ on the scale of the z-score-transformed data, and $\hat{\mu}_2$ is the variance of the ln-transformed distribution. Since we are using $\delta^{(z)}$ to calculate $\hat{\sigma}$, we must multiply by the standard deviation of the ln-transformed sample to put $\hat{\sigma}$ back into the original scale of measurement.

Given an estimate of σ, we can generate approximate confidence intervals for δ by setting $\sigma = \hat{\sigma}$, and generating samples with similar properties at different values of δ.

To continue the numeric example we return to the finger data from Table 1. To get our estimate of $\hat{\sigma}$, we substitute the previously derived value for $\hat{\delta}^{(z)} = 0.825$ and $\sqrt{\hat{\mu}_2} = 0.065$ into Equation 4, which gives us $\hat{\sigma} = 0.034$. Since we know the sex of these data, we can calculate the actual within-sex standard deviation. Since the numbers of males and females are slightly different, we must weight the average within-sex standard deviation accordingly. The average $\hat{\sigma}$ for males and females is 0.032, which is very close to the estimate made when sex is unknown.

To estimate 95% confidence intervals, we use a computer program to simulate sampling a distribution with a $\mu_1 = 2.232$, a $\sigma = 0.034$, and a sample size of 17. At each value of δ, starting at 0 and increasing by increments of 0.01, we simulated 500 samples. The smallest value of δ whose spread of $\hat{\delta}$ values includes our estimate of $\hat{\delta}$ is $\delta = 0.045$, and the largest value of δ whose spread includes our estimate is $\delta = 0.095$. In more conventional terms, the 95% confidence interval for our estimate is from $e^{2(0.045)}$ to $e^{2(0.095)}$, or 1.094, 1.209. The confidence interval for the estimate is large primarily because of the small sample size (n = 17) in this example.

RESULTS

The accuracy of the estimate of δ is sensitive to three factors: sample size n, the intrasexual standard deviation σ, and the population sex ratio. Our estimates are most accurate when: (1) sample size is large enough to accurately estimate the moments of the sample distribution; (2) male and female distributions overlap minimally; and (3) the sample sex ratio is approximately balanced. Significant deviations from these expectations tend to increase the sampling variance of $\hat{\delta}$.

Sample size

First, the sample size n strongly affects the spread of $\hat{\delta}$, with the spread becoming smaller as n increases from 50 to 100. Figure 2 shows the effect of increasing sample sizes with an approximately balanced sex ratio and $\sigma = 1$. For a given sample size, the dashed lines contain 95% of the 500 samples generated for each value of δ from 0 to 5 by increments of 0.5.

The MoM technique is a well behaved estimator, giving unbiased estimates of δ at all levels of dimorphism. $\hat{\delta}$ increases monotonically with larger values of δ. The distribution of $\hat{\delta}$ encloses and is closely centered on δ. The spread of $\hat{\delta}$ is smaller with larger sample sizes and smaller degrees of overlap between intrasexual distributions (i.e., when $\delta > \sigma$).

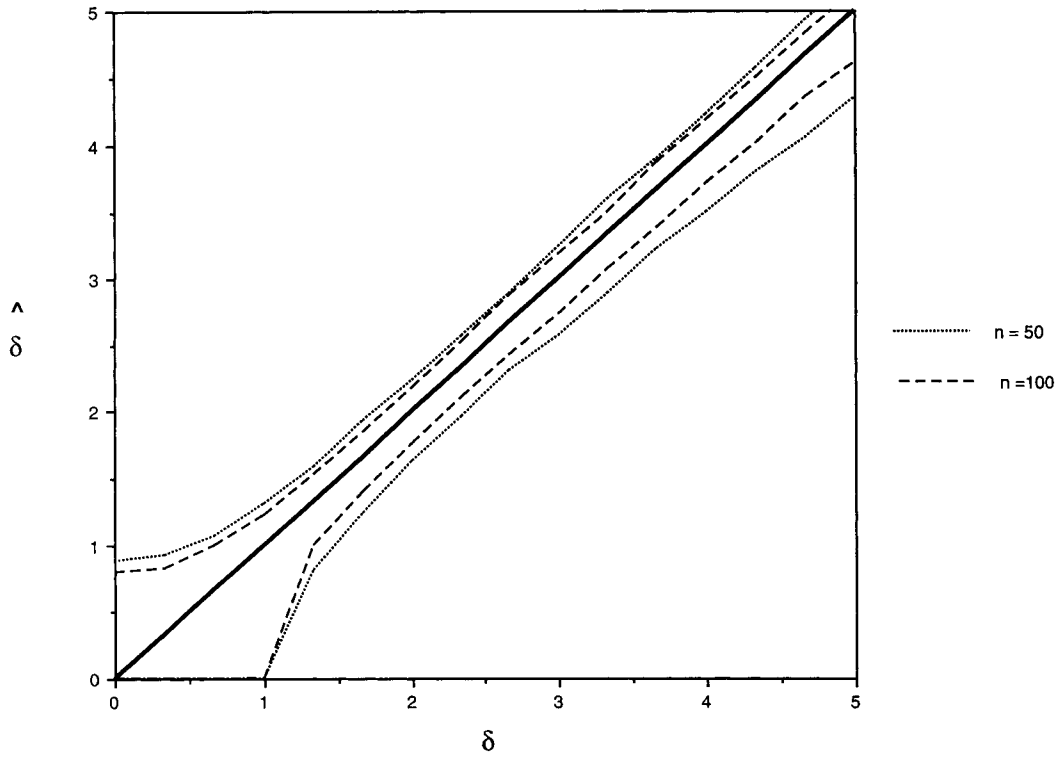

Fig. 2. The estimate of δ ($\hat{\delta}$) at different values of δ, with the within-sex standard deviation (σ) equal to 1, and an approximately balanced sex ratio. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

The method tends to underestimate δ slightly at small sample sizes (n = 50), which would result in a conservative estimate of sexual dimorphism. This tendency to underestimate δ in small sample sizes is a consequence of increasing variability both in $\hat{\sigma}$ and in the sample sex ratio. We are allowing $\hat{\sigma}$ and the sample sex ratio to vary, so random deviations produce larger effects in small samples than in large samples.

Within-sex variance

The reliability of our estimate is affected by the variance within the underlying male and female distributions. When δ is smaller than σ, the spread of $\hat{\delta}$ is wide, especially at small sample sizes (see Fig. 2). When $\delta \leq \sigma$, the resolution of the MoM technique is limited. This behavior is similar to that described by Godfrey et al. (1993), who demonstrated that low levels of dimorphism (male mean <28% larger than the female mean) can often not be reliably distinguished from no dimorphism using a different method of finite mixture analysis.

The MoM technique performs remarkably well even if its assumptions are violated. We assumed, for instance that the σ values for both sexes are equal, but this may not be correct (Gaulin and Sailer, 1984; Godfrey et al., 1993; Kay et al., 1988; Plavcan and Kay, 1988). If σ is larger in one sex than the other, this affects the estimate of δ. This effect can be seen in Figure 3, where the σ in males is 2.0 and the female σ is 1.0. $\hat{\delta}$ has greater sampling variance with larger values of σ, even if σ is larger in only one sex. This effect is the same regardless of which sex has larger values of σ. There is also a downward bias when δ/σ is small.

Biased sex ratios

Taphonomic processes may skew sex ratios by preferentially preserving the more

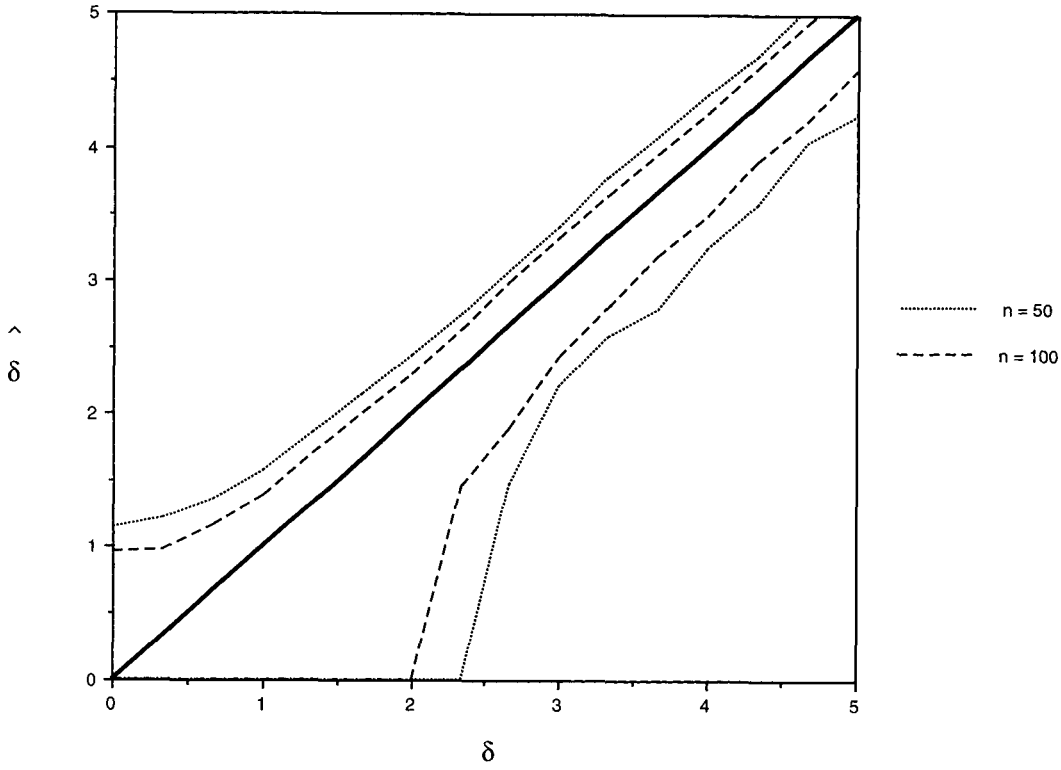

Fig. 3. The estimate of δ ($\hat{\delta}$) at different values of δ, with the male within-sex standard deviation (σ) equal to 2, and the female σ = 1, and an approximately balanced sex ratio. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

robust, dense (and likely male) specimens from an initial assemblage. Given this potential bias, it is important to know the effect of a deviation from a balanced sex ratio on the estimates of δ.

Skewed sex ratios do indeed affect the spread of $\hat{\delta}$ (Fig. 4). At all sample sizes, having a sex ratio of 60/40 (males/females) underestimate δ. This effect is the same whether the sample includes an excess of males or an excess of females. Differences in σ between sexes and biased sex ratios move $\hat{\delta}$ toward 0, again making the MoM technique a conservative estimator. The fact that we will not know the sex ratio in samples argues for using an estimator which is not inflated by skewed sex ratios and for using traits that are less sensitive to taphonomic skewing. Teeth, for example, are less sensitive to differential preservation than other skeletal elements (Lyman, 1985).

DISCUSSION

The MoM procedure has a number of advantages over previous methods. First, it allows us to put confidence intervals around the estimates of dimorphism. Previous techniques have allowed the accuracy of estimates of dimorphism to be assessed only in terms of roughly defined upper and lower limits (see Plavcan, 1994). Since the MoM technique defines meaningful 95% confidence intervals, we can reliably discriminate between different degrees of dimorphism, limited primarily by the size of paleontological samples.

Second, unlike previous techniques, the MoM requires no living species analogs. Analogs are living species (about which much is known) which are assumed to be sufficiently similar to fossil species that they can be used to make inferences about them. The problem

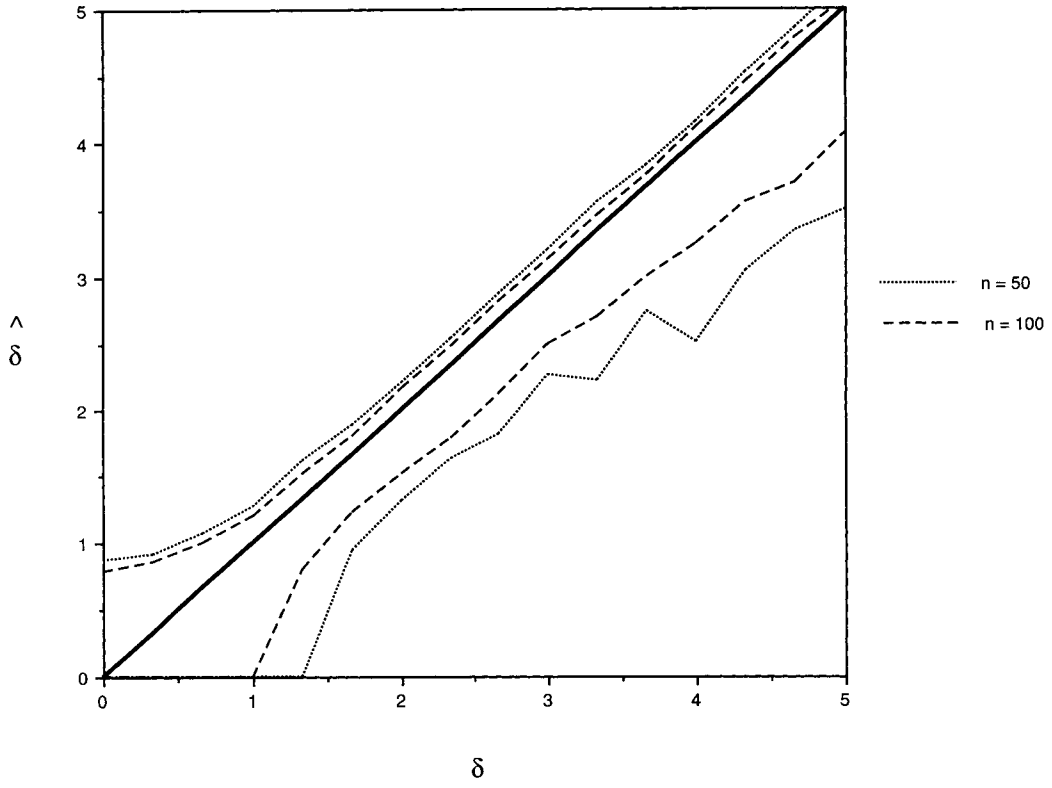

Fig. 4. The estimate of δ ($\hat{\delta}$) at different values of δ, with the within-sex standard deviation (σ) equal to 1, and a sex ratio of 60/40. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

is that we cannot be sure which living species are good analogs for fossil species. The need for analogs is not limited to the coefficient of variation technique (reviewed by Plavcan, 1994, below) but applies to all previous techniques. To produce confidence intervals for an estimate of dimorphism, we must first estimate the within-sex standard deviation σ. Previously σ has been estimated by using living organisms as analogs (Plavcan, 1994: 475). Since the MoM technique does not require analogs, we avoid the unknown error associated with using them.

It would be useful to know how well the MoM technique performs relative to other techniques. Recently, Pavcan (1994) reviewed several techniques for estimating dimorphism in fossil species. We use a procedure similar to Plavcan's and compare two previous techniques to the MoM technique.

Comparison of methods

We evaluated our method along with two tested by Plavcan: the coefficient of variation (CV) method (Fleagle et al., 1980; Kay, 1982a; Leutenegger and Shell, 1987) and the mean method (Godfrey et al., 1993; Plavcan, 1994). We used a hypothetical trait with a mean of 10.0 and, as in Plavcan's study, within-sex coefficients of variation (the standard deviation divided by the mean, multiplied by 100) of 14.0 and 5.5. We again allowed the sample sex ratio and $\hat{\sigma}$ to vary. We tested each technique with 500 samples of known dimorphism at each level from 1.0 (no dimorphism) to 2.2 (males >2 times as

large as females) by increments of 0.1. We defined dimorphism as the male mean divided by the female mean (M/F) for the mean and CV techniques, and $e^{2\delta}$ for the MoM technique.

Mean method. The simplest method used to determine dimorphism is the mean method. The combined-sex sample is simply divided at the mean. The large half of the distribution is assumed to represent males, and the small half females. The mean of the "male" subsample is divided by the mean for the "female" subsample to estimate dimorphism.

The mean method works poorly at low to moderate degrees of dimorphism, consistently overestimating dimorphism when the within-sex coefficient of variation is equal to 14.0 (Fig. 5a). It consistently overestimates dimorphism until the dimorphism rate is >1.8. There is less bias at higher levels of dimorphism but the spread of the estimate increases. Comparison of the sample sizes (n = 50 and n = 100) shows that the estimates improve somewhat with larger sample sizes, but the method provides reasonably accurate estimates only when the degree of sexual dimorphism is large.

This method performs better when the within-sex coefficient of variation is reduced to 5.5 (Fig. 5b). The region of over-estimation is reduced, providing accurate and unbiased estimates of dimorphism when the degree of dimorphism is greater than between 1.2 and 1.3.

Our conclusion is that the mean method works well only with clearly bimodal distributions, when the mean of the combined-sex sample more accurately discriminates the males from the females (see also Godfrey et al., 1993; discussion below). A major problem with this technique is that it can be difficult to accurately assess whether a distribution is actually bimodal or simply appears bimodal because of sampling. The need to rely on visual inspection of distributions limits the utility of this technique.

Coefficient of variation method. The coefficient of variation method relies on the fact that, as the means of the two underlying within-sex distributions get farther apart, the variation of the sample increases. This variation is usually expressed as a coefficient of variation (CV), which is the sample standard deviation divided by the sample mean. Using the CVs of species of known dimorphism to predict the dimorphism of unknown species requires a regression equation estimated from living species. While the reliability of this method has been criticized (Martin and Andrews, 1984; Vitzhum, 1990), the accuracy of this method under ideal conditions is of interest.

To generate the regression formula necessary to interpret the CVs, we generated 18 samples of size n = 50 for two levels of within-sex coefficient of variation, 14.0 and 5.5. These samples are equivalent to hypothetical "species" used to estimate the regression equations in applications of the CV method (Fleagle et al., 1980; Kay, 1982a; Leutenegger and Shell, 1987). Since the mean of a trait likely varies between analog species, we allowed the mean to vary between samples (between 7.5 and 12.5).

The hypothetical samples used to derive our regression equation are unrealistically good for several reasons. First, the number of "species" used to generate this regression equation is much larger than the number of analogs often used (see Kay, 1982a); Leutenegger and Shell, 1987), and contains "species" at all levels of dimorphism. Second, each of our species had the same within-sex variation, σ, and the same within-sex coefficient of variation. Had we allowed σ or the within-sex coefficient of variation to vary, the regression equation would have been less accurate. Third, the estimate of the mean and variance of the trait within each sample are accurate, as the sample size for each species is large (n = 50).

The relationship between CV and dimorphism has been argued to be nonlinear, so Plavcan ln-transformed his simulated species prior to deriving his regression equation (Leutenegger and Shell, 1987; see Plavcan, 1993:469). While this step is unnecessary with data that are already normal, we ln-transformed our measure of known dimorphism prior to deriving our regression equation in order to make our methods as similar

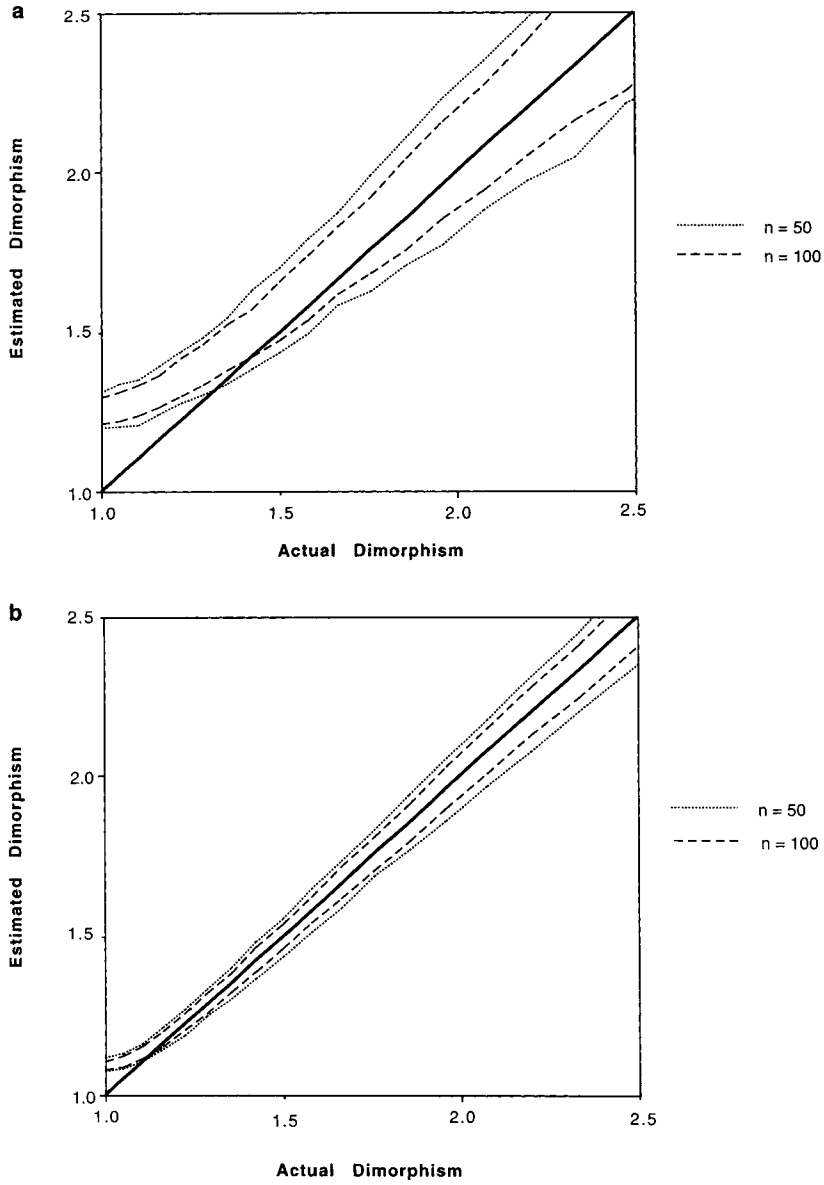

Fig. 5. The estimate of dimorphism at different values of actual dimorphism using the mean method, with an approximately balanced sex ratio, a hypothetical trait with a $\bar{x} = 10$, and a within-sex coefficient of variation of **(a)** 14.0 and **(b)** 5.5. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

to Plavcan's as possible (Fig. 6). The regression equation derived from our hypothetical species is (with a within-sex coefficient of variation = 14.0):

$$Y = -0.18712 + 1.9740\,X$$

where Y is ln(dimorphism) and X is the distribution CV. We used this regression equation to interpret the CVs of the 5,500 simulated samples.

The CV method does not perform well under these conditions (Fig. 7a). At low levels

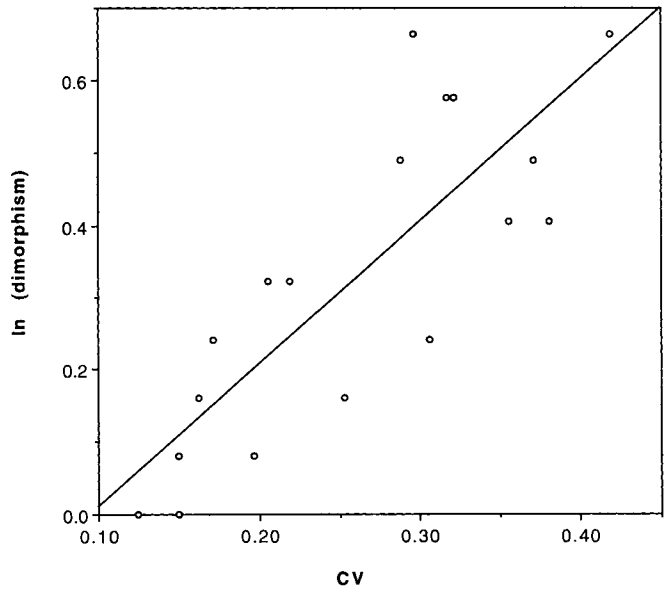

Fig. 6. Hypothetical "species" derived using the procedure described in the text, with $\bar{x}$ between 7.5 and 12.5, within-sex coefficient of variation of 14.0, and approximately balanced sex ratios. The regression equation for the line is $Y = -0.18712 + 1.9740\ X$, $r^2 = 0.673$, $n = 18$.

of dimorphism, the regression equation overestimates degree of dimorphism, a problem little ameliorated by larger samples. As the degree of dimorphism increases, so does the spread of the estimates of dimorphism. Decreasing the within-sex coefficient of variation to 5.5 (regression equation $Y = -0.024558 + 2.0433\ X$, $r^2 = 0.864$, $n = 18$) improves the resolution of this technique somewhat, but does not eliminate the biases observed earlier (Fig. 7b). This happens in part because of the way CVs are calculated. The mean, for example, affects the value for the coefficient of variation and yet in itself contains no information about dimorphism. There is no a priori reason why a trait with a mean of 10 (in whatever units) should be more or less dimorphic than a trait with a mean of 100.

It is difficult to estimate the within-sex coefficients of variation in fossil species, particularly if samples are small and living analogs are few. Variation in the within-sex coefficients of variation between living species can make it difficult to know which species are the appropriate analogs. Given that there are often few living species available from which to derive a regression equation (see Leutenegger and Shell, 1987), the accuracy of this method may be subject to error which is difficult or impossible to estimate.

Our conclusion is that the CV method generates imprecise and often biased estimates of sexual dimorphism, even when the samples used to generate the regression equation are unrealistically numerous and critical information is known with some accuracy. Larger samples do not substantially improve the resolution of this technique.

Method-of-moments technique. Figure 8a shows the results of the MoM technique with a within-sex coefficient of variation of 14.0. With this within-sex coefficient of variation, the technique performs best when the degree of dimorphism is greater than 1.3. Below this, δ is smaller than σ and the resolution is insufficient to reliably distinguish smaller degrees of dimorphism from no dimorphism at all. This can also be seen when the within-sex coefficient of variation is reduced to 5.5 (Fig. 8b). δ becomes larger than σ when the degree of dimorphism is approximately 1.15, after which the resolution of the estimate improves. At all levels of dimor-

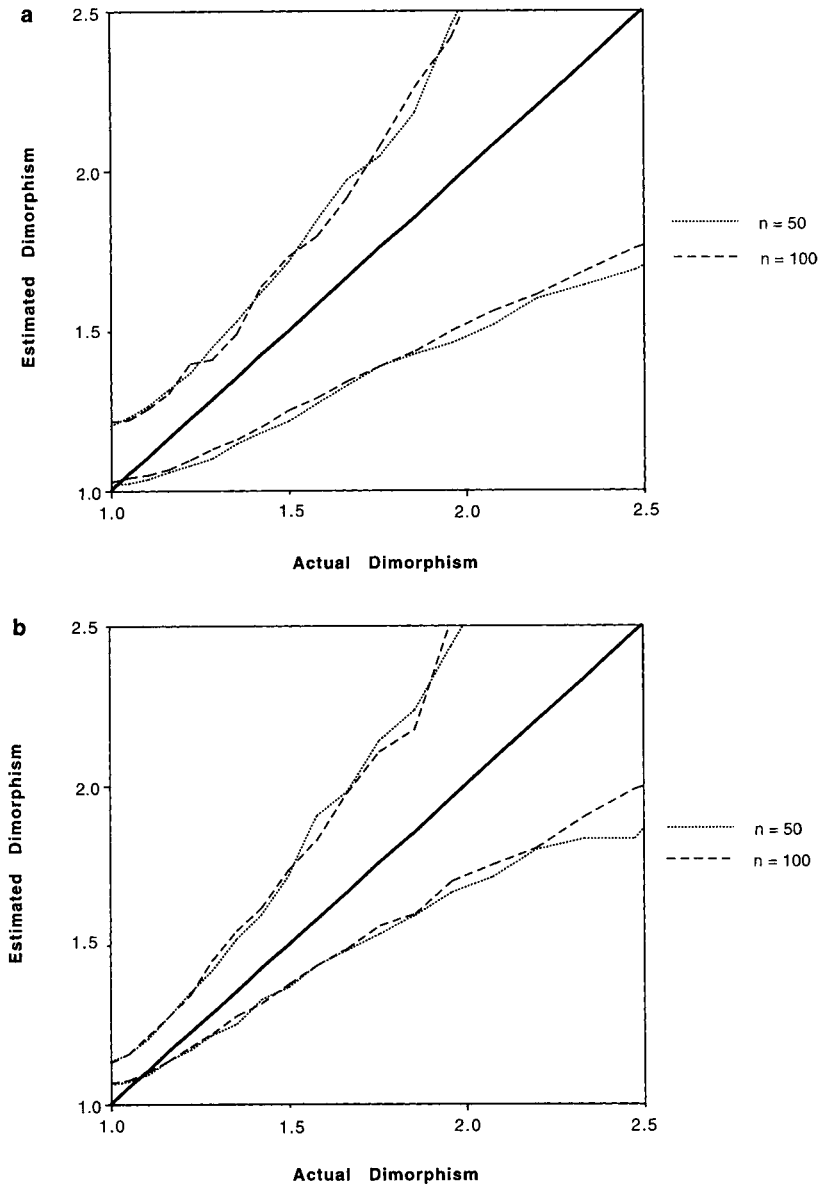

Fig. 7. The estimate of dimorphism at different values of actual dimorphism using the CV method, with an approximately balanced sex ratio, a hypothetical trait with a $\bar{x} = 10$, and a within-sex coefficient of variation of **(a)** 14.0 and **(b)** 5.5. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

phism, the estimates of dimorphism enclose the actual values of dimorphism, showing that the MoM technique is an unbiased estimator. Larger sample sizes greatly improve the resolution of this technique.

CONCLUSIONS

The method-of-moments technique has several features to recommend it over previous techniques. While the method works best when the overlap between the male and

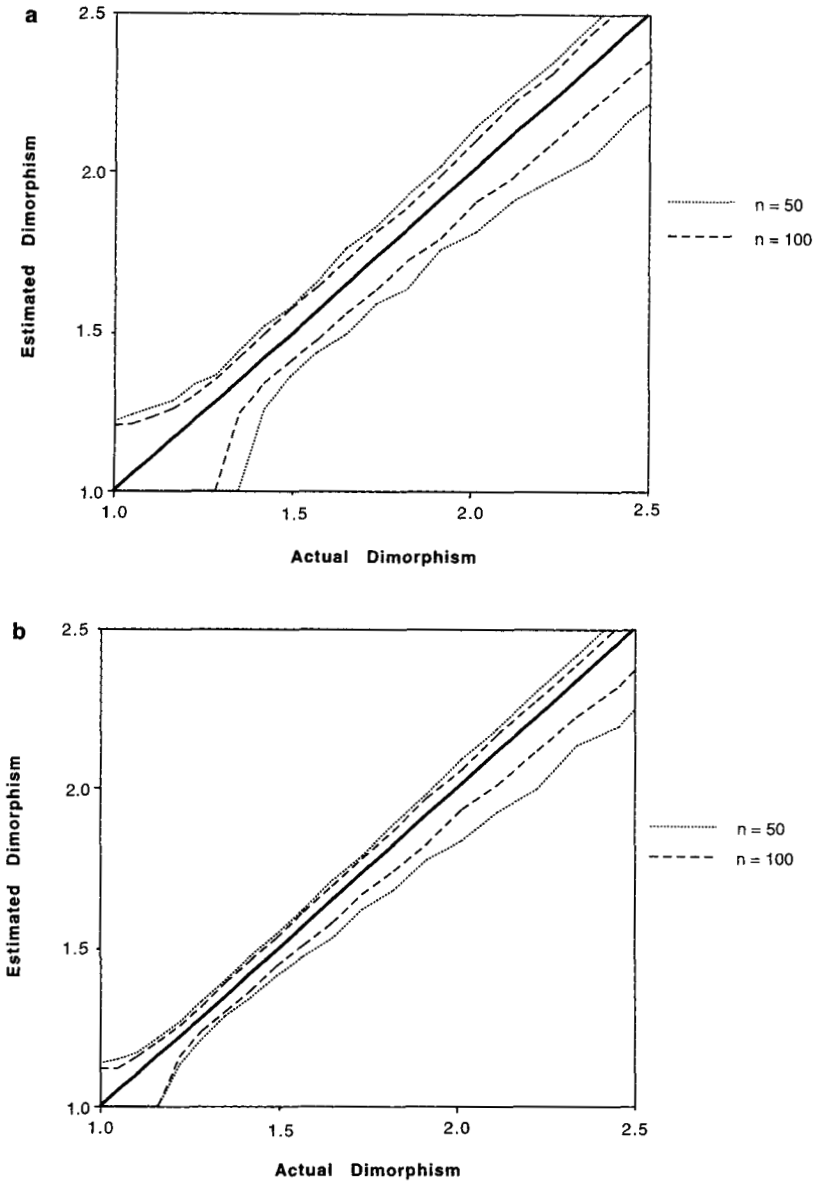

Fig. 8. The estimate of dimorphism at different values of actual dimorphism using the method-of-moments, with an approximately balanced sex ratio, a hypothetical trait with a $\bar{x} = 10$, and a within-sex coefficient of variation of **(a)** 14.0 and **(b)** 5.5. The dashed lines enclose the central 95% of 500 simulated samples at different sample sizes.

female distributions is minimal, it performs well even when overlap is substantial. It provides conservative and unbiased estimates at all levels of dimorphism and allows us to generate 95% confidence intervals around these estimates. These estimates are made without reference to living species analogs or visual evaluations of bimodality, and so avoids the often unknowable error asociated with these procedures. This technique also

provides us with an opportunity to reliably measure dimorphism in different traits in the same species.

ACKNOWLEDGMENTS

We thank Tessa Nang, Dennis O'Rourke, Mike Plavcan, Mary Stiner, Joan Coltrain, Duncan Metcalfe, Kristen Hawkes, Doug Edwards, Melissa Heck, Matt Wimmer, Lori Hunsacker, Steve Leigh, and Lyle Konigsberg for help and comments. We dedicate the paper to the memory of Elizabeth Janine Manderscheid.

LITERATURE CITED

Andersson M (1994) Sexual selection. Princeton, NJ: Princeton University Press.

Andrews P (1983) The natural history of *Sivapithecus*. In RL Ciochon and RS Corruccine (eds.): New Interpretations of Ape and Human Ancestry. New York: Plenum, pp. 441–463.

Clutton-Brock TH, and Harvey PH (1977) Primate ecology and social organization. J. Zool. Lond. *183:*1–39.

Clutton-Brock TH, and Harvey PH (1978) Mammals, resources and reproductive strategies. Nature (Lond.) *273:*191–195.

Cope DA (1993) Measures of dental variation as indicators of multiple taxa in samples of sympatric *Cercopithecus* species. In WH Kimbel and LB Martin (eds.): Species, Species Concepts, and Primate Evolution. New York: Plenum, pp. 211–237.

Everett BS, and Hand DJ (1981) Finite Mixture Distributions. New York: Chapman and Hall.

Fleagle JG, Kay RF, and Simons EL (1980) Sexual dimorphism in early anthropoids. Nature *287:*328–330.

Foley RA, and Lee PC (1989) Finite social space, evolutionary pathways, and reconstructing hominid behaviors. Science *243:*901–243.

Gaulin SJC, and Sailer LD (1984) Sexual dimorphism in weight among the Primates: The relative impact of allometry and sexual selection. Int. J. Primatol. *5:*515–535.

Gingerich PD (1981) Cranial morphology and adaptations in Eocene Adapidae I. Sexual dimorphism in *Adapis mangnus* and *Adapis parisiensis*. Am. J. Phys. Anthropol. *56:*217–234.

Godfrey LR, Lyon SK, and Sutherland MR (1993) Sexual dimorphism in large-bodied primates: The case of the subfossil lemurs. Am. J. Phys. Anthropol. *90:*315–334.

Groves GP (1989) A theory of human and primate evolution. Oxford: Clarendon Press.

Harvey PH, Kavanagh M, and Clutton-Brock TH (1978) Sexual dimorphism in primate teeth. J. Zool. Lond. *186:*475–485.

Hasting NAJ, and Peacock JB (1974) Statistical distributions: A handbook for students and practitioners. New York: Wiley.

Jarman P (1983) Mating system and sexual dimorphism in large, terrestrial, mammalian herbivores. Biol. Rev. Camb. Philos. Soc. *58:*485–520.

Johanson DC, and White TD (1979) A systematic assessment of early African hominids. Science *202:* 321–330.

Johanson DC, White TD, and Coppens Y (1983) Dental remains from the Hadar Formation, Ethiopia: 1974–1977 collections. Am. J. Phys. Anthropol. *57:*545–603.

Johnson NL, and Kotz S (1970) Continuous univariate distributions. New York: Houghton Mifflin.

Jungers WL (1990) Problems in reconstructing body size in fossil primates. In J Damuth and BJ MacFadden (eds.): Body Size in Mammalian Paleobiology. Cambridge: Cambridge University Press, pp. 103–118.

Kay RF (1982a) Sexual dimorphism in the Ramapithecinae. Proc. Natl. Acad. Sci USA *79:*209–212.

Kay RF (1982b) *Sivapithecus simonsi,* a new species of Miocene hominoid, with comments on the phylogenetic status of the Ramapithecinae. Int. J. Primatol. *3:*113–173.

Kay RF, Plavcan JM, Glander KE, and Wright PC (1988) Sexual selection and canine dimorphism in New World monkeys. Am. J. Phys. Anthropol. *77:*385–397.

Kelley J (1993) Taxonomic implications of sexual dimorphism in *Lufengpithecus*. In WH Kimbel and LB Martin (eds.): Species, Species Concepts, and Primate Evolution. New York: Plenum, pp. 429–458.

Kelley J, and Xu Q (1991) Extreme sexual dimorphism in a Miocene hominoid. Nature *352:*151–153.

Kimbel WH, and White TD (1988) Variation, sexual dimorphism and the taxonomy of *Australopithecus*. In FE Grine (ed.): Evolutionary History of the "Robust" Australopithecenes. New York: Aldine de Gruyter, pp. 175–192.

Krishtalka L, Stucky RK, and Beard KC (1990) The earliest fossil evidence for sexual dimorphism in primates. Proc. Natl. Acad. Sci. USA *87:*5223–5226.

Leutenegger W, and Shell B (1987) Variability and sexual dimorphism in canine size of *Australopithecus* and extant hominoids. J. Hum. Evol. *16:*359–367.

Lieberman DE, Pilbeam DR, and Wood BA (1988) A probabilistic approach to the problem and sexual dimorphism in *Homo habilis:* A comparison of KNM-ER 1470 and KNM-ER 1813. J. Hum. Evol. *17:* 503–511.

Lyman RL (1985) Bone frequencies: Differential destruction, in situ destruction, and bone transport and their archaeological implications. J. Archaeol. Scien. *12:* 221–236.

Martin L (1991) Teeth, sex and species. Nature (Lond.) *352:*111–112.

Martin LB, and Andrews P (1984) The phylogenetic position of *Graecopithecus freyburgi* Koenigswald. Cour. Forschungsinst Sencken b. *69:*25–40.

McHenry HM (1991) Sexual dimorphism in *Australopithecus afarensis*. J. Hum. Evol. *20:*21–32.

Miller JA (1991) Does brain size variability provide evidence of multiple species in *Homo habilis?* Am. J. Phys. Anthropol. *84:*385–398.

Oxnard CE (1983) Commentary on two misconceptions of phylogeny and classification Am. J. Phys. Anthropol. *59:*133–170.

Pickford M (1986) Sexual dimorphism in Proconsul. In M Pickford and B Chiarelli (eds.): Sexual Dimorphism

in Living and Fossil Primates. Florence, Italy: Il Sedicesimo, pp. 133–170.
Plavcan JM (1993) Catarrhine dental variability and species recognition in the fossil record. In WH Kimbel and LB Martin (eds.): Species, Species Concepts, and Primate Evolution. New York: Plenum, pp. 239–263.
Plavcan JM (1994) Comparison of four simple methods for estimating sexual dimorphism in fossils. Am. J. Phys. Anthropol. *94:*465–476.
Pickford M, and Chiarelli B (eds.) (1986) Sexual dimorphism in living and fossil primates. Florence, Italy: Il Sedicesimo, pp. 133–170.
Plavcan JM, and Kay RF (1988) Sexual dimorphism and dental variability in Platyrrhine Primates. Int. J. Primatol. *9:*169–178.
Plavcan JM, and van Schaik CP (1992a) Intrasexual competition and canine dimorphism in anthropoid primates. Am. J. Phys. Anthropol. *87:*461–477.
Plavcan JM, and van Schaik CP (1992b) Female canine size in anthropoid primates. Am. J. Phys. Anthropol. Suppl. *14:*132.
Ralls K (1977) Sexual dimorphism in mammals: Avian models and unanswered questions. Am. Nat. *111:* 917–938.
Richard AF (1992) Aggressive competition between males, female-controlled polygyny and sexual monomorphism in a Malagasy primate, *Propithecus verreauxi.* J. Hum. Evol. *22:*395–406.
Rightmire GP (1993) Variation among early *Homo* crania from Olduvai Gorge and the Koobi Fora region. Am. J. Phys. Anthropol. *90:*1–33.
Stern JT, and Susman RL (1983) The locomotor anatomy of *Australopithecus afarensis.* Am. J. Phys. Anthropol. *60:*279–317.
Teaford MF, Walker A, and Mugaisi GS (1993) Species discrimination in Proconsul from Rusinga and Mfangano islands, Kenya. In WH Kimbel and LB Martin (eds.): Species, Species Concepts, and Primate Evolution. New York: Plenum, pp. 373–392.
Tobias PV (1978) The earliest Transvaal members of the genus *Homo* and another look at some problems of hominid taxonomy and systematics. Z. Morphol. Anthropol. *69:*225–265.
Tobias PV (1991) Olduvai Gorge, Volume 4. The Skulls, Endocasts and Teeth of *Homo habilis.* Cambridge: Cambridge University Press.
Vitzhum VJ (1990) Odontometric variation within and between taxonomic levels of Cercopithecidae: Implications for the interpretations of fossil samples. Hum. Evol. *5:*359–374.
Walker AC (1981) The Koobi Fora hominids and their bearings on the origins of the genus *Homo.* In BA Sigmon and JS Cybulski (eds): *Homo erectus.* Papers in Honor of Davidson Black. Toronto: University of Toronto Press, pp. 193–215.
Wood BA (1985) Sexual dimorphism in the hominid fossil record. In J Ghesquirer, RD Martin, and F Newcombe (eds.): Human Sexual Dimorphism. London: Taylor and Francis, pp. 105–123.
Wood BA (1991) Koobi Fora Research Project, Volume 4. Hominid Cranial Remains. Oxford: Clarendon Press.
Wood BA (1992) Origin and evolution of the genus *Homo.* Nature *355:*783–790.
Zihlman AL (1985) *Australopithecus afarensis:* Two sexes or two species? In PV Tobias (ed.): Hominid Evolution: Past and Future. New York: Alan R. Liss, pp. 213–220.

APPENDIX

Equations 1 and 2 were derived using the following steps, equating the measured moments about the origin $\hat{m}_i$ with the theoretical moments about the origin m_i using equations from Johnson and Kotz (1970: 89).

m_1 is equivalent to the mean of the combined-sex distribution, and is the average of the means of the two underlying distributions. To simplify exposition, we refer to the mean of the larger distribution as M for "males" and the mean of the smaller distribution as F for "females." m_1 is the sum of the two underlying means:

$$m_1 = w_1M + w_2F \tag{5}$$

where w_1 = the proportion of males and w_2 = the proportion of females. We assume that any particular specimen has equal probability of being a male or a female, so $w_1 = w_2 = 1/2$. M and F are therefore equally distant from m_1, a distance we call δ (Fig. 1a). M thus can be expressed as $M = m_1 + \delta$, and the female mean as $F = m_1 - \delta$.

m_2 is equivalent to the variance of a distribution, and can be expressed in terms of the underlying distributions. To simplify this derivation, we assume that the variances of the underlying male and female distributions are equal, so both are simply σ^2. Substituting the aforementioned values for σ^2, w, M, and F into the equation for m_2 from Johnson and Kotz (1970: 89) produces:

$$m_2 = \tfrac{1}{2}(m_1 + \delta)^2 + \sigma^2 + \tfrac{1}{2}(m_1 - \delta)^2. \tag{6}$$

Solving Equation 6 for σ^2 yields Equation 1.

m_3 measures symmetry about the mean. Distributions where m_3 deviates from 0 have more probability mass on one side of the mean than on the other. Since we assume that the two underlying distributions are normal and symmetric, the combined-sex distribution is symmetric as well. The third

moment of the combined-sex distribution therefore provides no information about dimorphism.

m_4 is sensitive to the "flatness" or "peakedness" in distributions, and is equal to (Johnson and Kotz, 1970: 89):

$$m_4 = w(M^4 + 6M^2\sigma^2 + 3\sigma^4) + (1 - w)(F^4 + 6F^2\sigma^2 + 3\sigma^4). \quad (7)$$

Substituting aforementioned values for w, M, F, and σ^2 and solving for δ yields Equation 2.